#23

CERTIFICATE OF DELIVERY			
I hereby certify that this correspondence is being hand-delivered to the U.S. Patent and Trademark Office, Art Unit 1655			
Typed or Printed Name	Doanh Vu		
Signature	Doanh Kim Vu	**Date**	7-5-02

SUPPLEMENTAL DECLARATION UNDER 37 C.F.R. § 1.132 Address to: Assistant Commissioner for Patents Washington, D.C. 20231		
	Attorney Docket Confirmation No.	CLON-015
	First Named Inventor	Chenchik et al.
	Application Number	09/440,829
	Filing Date	November 15, 1999
	Group Art Unit	~~1655~~ 1634
	Examiner Name	Forman, B.
	Title	*Long Oligonucleotide Arrays*

Sir:

I, Alex Chenchik, am a co-inventor of the above referenced application and an employee of Clontech Laboratories, Inc., the assignee of the above the above referenced application. A copy of my C.V. is already of record in this application.

I hereby declare the following:

I. Summary of Experimental Section of the Above Captioned Application

A. Test Array

As explained in my previous declaration, in the working exemplification of the above-captioned application, an array was prepared in which the probe lengths ranged from 50 to 100 nt. See Examples 2 to 4 of the above captioned application. In the prepared array, each spot was made up of equal length probes, such that one spot had only 50 nt length probes, one spot had only 60 nt length probes, one spot had only 70 nt length probes, one spot had only 80 nt length probes, one spot had only 90 nt length probes and one spot had only 100 nt length probes. The specific probe sequences displayed on the array are provided on pages 41 and 42 of the specification, in Table 1.

B. Generation of labeled target.

A radio-isotopically labeled population of target nucleic acids that included a target nucleic acid for each probe on the array described above present in an equimolar amount was then prepared according to the protocol provided in Example 1. Thus, the prepared population of target nucleic acids included an equimolar amount of each different target nucleic acid member in the population. Furthermore, the each different probe on the array had a fully complementary target nucleic acid in the preparation population of labeled target nucleic acids.

C. Hybridization

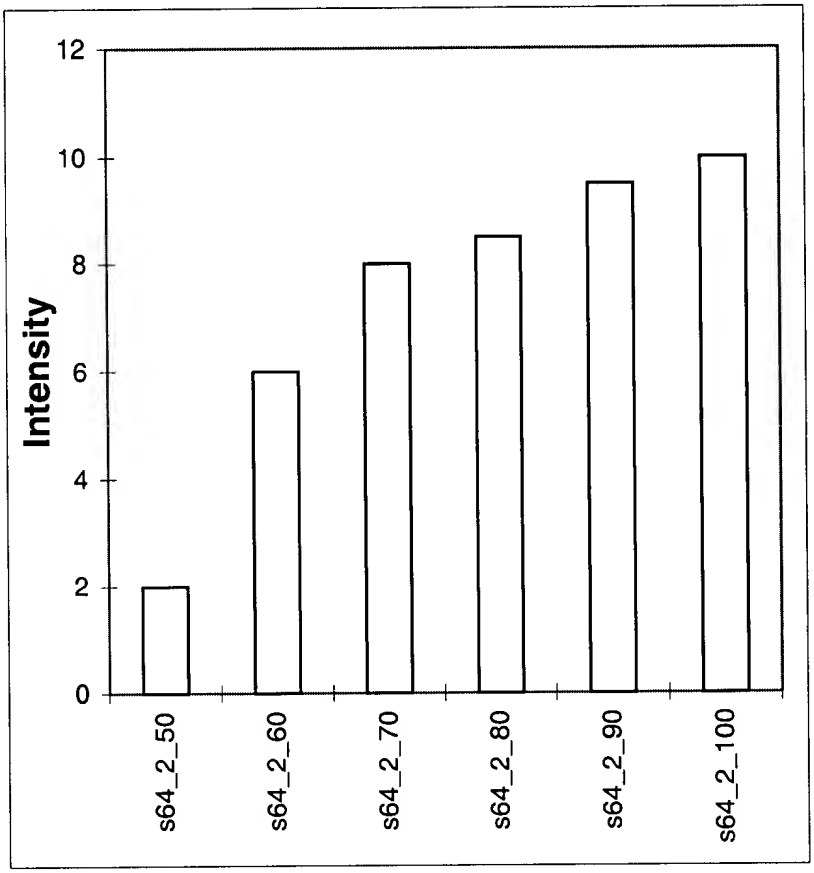
The population of labeled target nucleic acids was then hybridized to the array using the protocol described in Example 5. Briefly, the population of target nucleic acids was contacted with the array under hybridization conditions and maintained for a set period of time, followed by washing to stop hybridization. Thus, each target was in contact with the array under the same hybridization conditions for the same period of time.

Any resultant hybrids present on the array were then detected by phosphorimager.

D. Evaluation of Hybridization Efficiency

Hybridization efficiency is a term well known to those of skill in the art. The term describes efficiency of interaction between two complementary nucleic acids. In other words, the term describes the yield of hybrids between target and probe after a hybridization based assay. "Intensity of hybridization signal" is an experimental measurement that corresponds directly to the yield of hybrid formation or hybridization efficiency and thus is used as a measure of hybridization efficiency. **Thus, hybridization efficiency is measured by looking at the intensity of hybridization signal.** As such, measuring relative intensities of the signals corresponding to different size oligonucleotide probes provides a measurement of the relative efficiency of hybridization between target and probe.

Accordingly, as described in Example 6, signal intensity was measured using a phosphorimager and the results were employed to prepare the following figure:

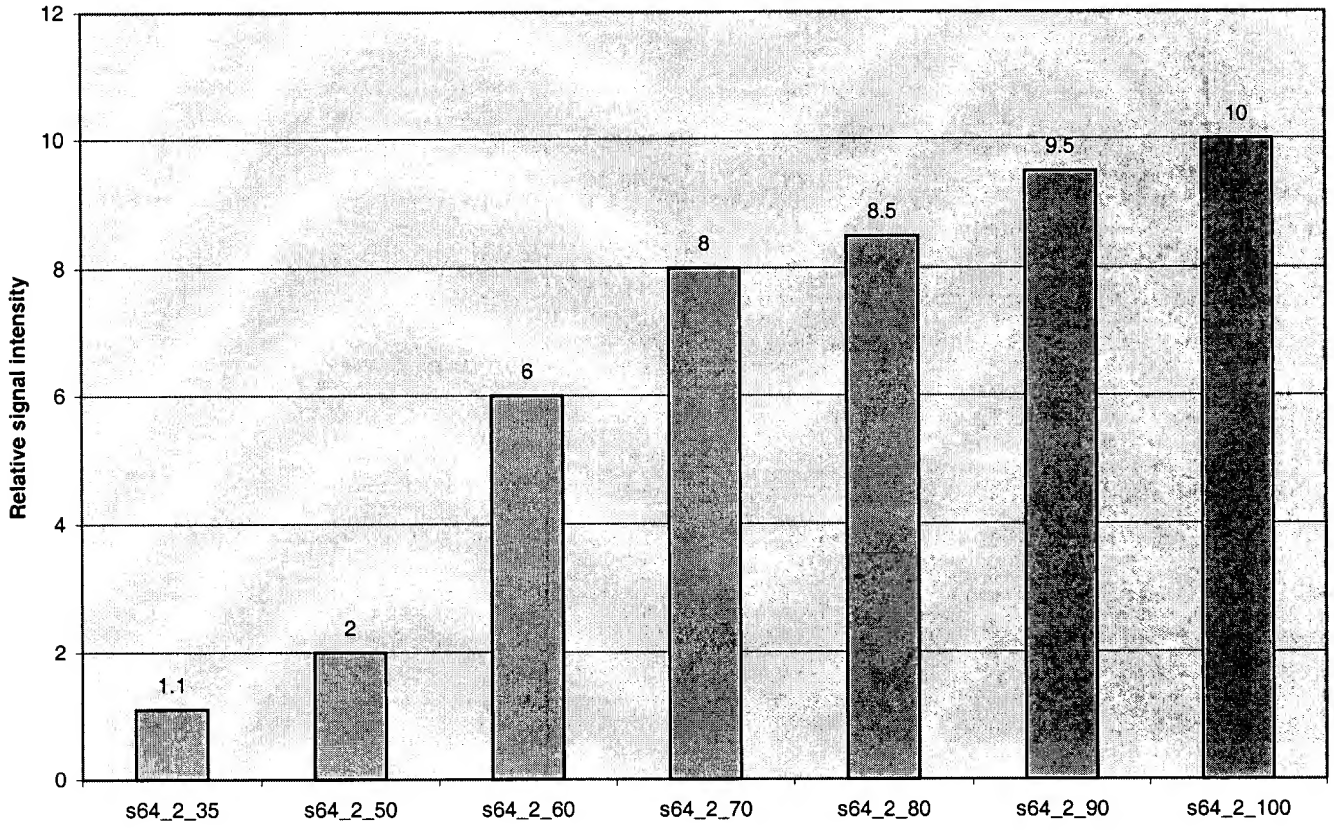


Because signal intensity is directly proportional to the hybridization efficiency, as described above, Figure 1 shows the hybridization efficiency of each of the different probe lengths on the array.

Figure 1 also shows that the hybridization efficiency dramatically depends on the size of the immobilized probe in the range between 50 and 100 nucleotides. For example differences in

the hybridization yield between 50 mer and 60, 80, 100-mer is about 3, 4 and 5-fold respectively higher for the longer probes. This finding means that hybridization efficiency is proportional to the 5-10 power of the length of the probe in this size range.

Essentially the same experiment was performed with the addition of 35-mer length probes. The results of this experiment are provided graphically below:



As can be seen from the above plot, the hybridization efficiency (as provided in terms of signal intensity) unexpectedly increases over 35 to 100 nt length range. We also discovered that differences in hybridization efficiency demonstrated slow exponential growth (close to linear dependance from the target length) between 35 and 50 nt targets, dramatic exponential growth between 50 and 70 nt targets (proportional to the 5-10 power of the target length) and followed by close to linear growth between 70 and 100 nt.

II. Results That Were Expected

Prior to conducting the experiments reported in the specification and summarized above, it was my expectation, which is the same as that which those of skill in the art would expect, that the hybridization efficiency would not be what was actually observed, but would instead be one of the following:

1. Based on the teachings of hybridization chemistry in "Nucleic acid hybridization - a practical approach" (ed. by B.D. Hames & S.J. Higgins) 1990, IRL Press, pp. 62 and 77; a copy of which is enclosed as Exhibit A) it was found experimentally for solution phase hybridization in which both the target and probe are present in solution that the rate of re-annealing (hybridization) is proportional to square root of the length of the shorter strand. This finding means that if one compares the rate of hybridization of a 50-mer and a 100-mer, the 100-mer should hybridize at about a 1.4-fold higher rate (and efficiency) than 50-mer, so that there is less than a 2-fold difference in hybridization efficiency between the different 50-mer and 100-mer sizes.

With respect to solid phase-based hybridization assay with which the claimed arrays are employed where the probe is attached to the solid surface, the rate-limiting step for this type of assay can be different and not even depend on the size of the probe or the target. Nevertheless it was found that for diffusion-limited array hybridization (probe immobilized on array is the excess in comparison with target in the solution), the rate of hybridization (and intensity of the signal) was **inversely** proportional to the molecular weight of the target. From a practical point of view this finding means that the longer the target the less hybridization rate. For nucleation-limited reaction, the rate of hybridization was found independent of the molecular weight of the target. While no data was published for dependance of hybridization rate from probe length in the range 30 to 100-mers, based on the above teaching one of skill in the art would expect that the hybridization rate would not depend on probe length or would be inversely proportional to the length of the probe.

2. Based on the teaching of "DNA probes" (ed. by George H. Keller & Mark M. Manak) 1993, Stockton Press, pp.6-8; this reference teaches that the rate of solid-phase hybridization is **inversely proportional** to the size of the target, i.e. the longer target the less hybridization rate. As such, hybridization efficiency should decrease as probe length increases. A copy of this reference is enclosed as Exhibit B.

3. Based on the teaching of "DNA Probes: Applications of the principles of Nucleic acid hybridization" by James G. Wetmur in Critical Reviews in Biochem. and Mol. Biol. (1991) 26 (3/4):227-259; the author reviews the current knowledge of hybridization-based assay and states that while solid-phase hybridization is poorly studied so far and hybridization kinetics could be more complicated, the hybridization kinetics observed with array formats could be described as for solution phase based hybridization.

4. Conclusion:

As demonstrated above, prior to my work in this area as embodied in the present application and claims, the prior art teachings indicated that the expected hybridization rate should be **inversely** proportional for solid phase hybridization, such as array hybridization, or at best proportional to the square root of the length of the target if the hybridization efficiency followed solution phase hybridization efficiency findings.

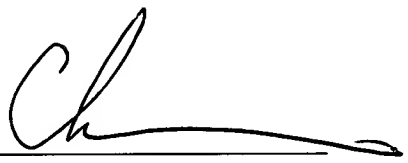
III. Conclusion

The present application and claims are based on our discovery of the unexpected, unusual behavior of the oligonucleotides in the range of 50 to 100-mers. The unexpected finding is that the hybridization rate (hybridization efficiency) dramatically depends on the size of the immobilized probe in the range between 50 and 100 nucleotides. For example differences in the hybridization yield between 50 mer and 60, 80, 100-mer is about 3, 4 and 5-fold respectively higher for the longer probes. This finding means that hybridization efficiency is proportional to the 5-10 power of the length of the probe in this size range. **This dependence is completely**

unpredictable based on the prior art knowledge, as summarized above.

As such, using a probe length of 50 to 100 nt provides unexpected results.

I further declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code; and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Date: July 1, 2002 Signature: 

Alex Chenchik

enc

- Exhibits A, B & C

Exhibit A

)ACH

Essex
3SQ, UK

of Leeds

uters in biology
'ia
:y testing
iistry
l and
luence analysis
l hybridisation
itide synthesis
esis:
isduction
ilture

ins
substances
:ometry
fluorimetry
nones
omas
nic stem cells
and translation

Nucleic acid hybridisation

a practical approach

Edited by

B D Hames

Department of Biochemistry,
University of Leeds, Leeds, England

S J Higgins

Department of Biochemistry,
University of Leeds, Leeds, England

IRL PRESS
Oxford · Washington DC

IRL Press
Eynsham
Oxford
England

First Published 1985
Reprinted 1986, 1987 (twice), 1988, 1990

British Library Cataloguing in Publication Data

Nucleic acid hybridisation : a practical approach.—(Practical approach series)
1. Nucleic acid hybridisation
I. Hames, B.D. II. Higgins, S.J. III. Series
547.7'9 QP620

ISBN 0-947946-61-6 (Hardbound)
ISBN 0-947946-23-3 (Softbound)

Cover illustration. On the left is an autoradiograph of λgt11 recombinants containing *Dictyostelium discoideum* genomic DNA inserts, screened on nitrocellulose filters with a nick-translated 4.1 kb repetitive genomic fragment by the method of Benton and Davis (see Chapter 5). On the right is an autoradiograph of part of *Drosophila melanogaster* polytene chromosome 2R (stained with Giemsa) showing multiple sites of *in situ* hybridisation by the mobile element pDmI 137 [Dawid *et al.* (1981) Cell **25**, 399]; magnification 1300 x. The photographs were kindly supplied by Ms. P. Jagger and Mr. D.P. Ramji (Department of Biochemistry, University of Leeds, UK) and Dr. M.L. Pardue (M.I.T., Cambridge, USA), respectively.

Printed by Information Press Ltd, Oxford, England.

cDNA. The theoretical curve which is predicted by Equations 20 – 22 is superimposed on the data, assuming that $k = 182$ litre. mol^{-1}. sec^{-1}. It is clear from *Figure 6* that the slowest reaction, which has a $D_0t_{½}$ of 5.5×10^{-3} mol. $litre^{-1}$. sec, occurs for $R_0 = D_0$. Thus if all reactions are taken to a D_0t value of at least 20 times this value, we can be sure that all reactions are complete irrespective of the ratio R_0/D_0. Hence, provided the rate constant of the reaction between the cDNA and its complementary mRNA is known, conditions can be chosen such that all reactions between a cDNA and an RNA whose complementary sequence content is unknown are complete.

6. FACTORS AFFECTING REACTION RATES

The rate at which complementary strands of nucleic acid form stable base-paired duplexes is dependent on a number of factors. These can broadly be divided into factors concerning the state of the nucleic acids, the conditions of incubation and the methods of detection used. If a set of reannealing or hybridisation experiments is carried out under identical conditions using nucleic acids of identical composition then the rates of reactions may be compared directly. If, however, as is more often the case, there are differences in the condition or composition of the nucleic acids, it is important to know which correction factors should be applied before reaction rates may be compared. The influence of such factors has been extensively reviewed elsewhere (16 – 18) and so only the main conclusions are presented here. Additional discussion of some of these factors is given in Chapter 4.

6.1 **Nucleic Acid Length**

The theoretical considerations of Wetmur and Davidson (1) predicted that the rate of reannealing of DNA would be proportional to the length (L) of the fragments. However it was found experimentally to be proportional to the square root of the length. This difference was attributed to an excluded volume effect which results in the complementary strands being unable to interact completely. The dependence on L was shown to hold over three orders of magnitude starting with L = 100 nucleotides.

In most DNA reannealing experiments the average length of the complementary strands is the same since DNA is usually randomly sheared prior to annealing. In some experiments, however, DNA strands of different lengths may be reannealed and it has been shown that the rate of such a reaction is proportional to the square root of the shorter of the annealing strands (19).

Hutton and Wetmur (20) studied rates of hybridisation using ϕX174 DNA and RNA transcribed with *E. coli* polymerase. Their results established clearly that the rate of hybridisation was proportional to the square root of length of the RNA, which in this case was the shorter strand. These results hold over a 5-fold range in molecular weight. Hence it has been well established that the rate of both DNA reannealing and RNA-DNA hybridisation is proportional to the square root of the length of the shorter strand.

6.2 **Base Composition**

Hutton and Wetmur (20) demonstrated that if the genome size and molecular weight of sheared DNA is taken into account, as indicated by Wetmur and Davidson (1), bacteriophage T2 DNA (34% G + C) and bacteriophage T7 DNA (50% G + C) have

in probe excess to being in filter-bound excess where, as we have seen, the kinetics are different. Increasing the concentration of probe in solution, $[C_s]$, will increase the initial rate of hybridisation at the filter and the proportion of filter-bound sequences in duplex will increase, but not dramatically. For DNA probes and filter-bound RNA, as in RNA dot blots, high concentrations of formamide can be used to suppress reassociation in solution (see Section 4.1.2).

3.1.2 *Single-stranded Probe*

Whether in excess or not, there is no reassociation of a single-stranded probe in solution unless there are regions of extensive self-complementarity. The rate of hybridisation to the filter and the amount of hybrid formed should increase with increase in $[C_s]$. It is important to note, however, that the probe concentration should not be increased without limit. If more than about 100 ng ^{32}P-labelled probe per ml is used, non-specific irreversible binding to the filter occurs.

3.2 **Probe Complexity**

For solution hybridisation, the rate of reassociation of DNA is an inverse function of its complexity, so that the more complex the DNA, the slower the rate of reassociation (2,12). Extending this to filter hybridisations, the rate of reassociation of the probe should fall when the complexity of the DNA increases and its effective $[C_s]$ decreases. This is indeed what is observed (9). In contrast, two effects of complexity are seen for hybridisation of the probe to filter-bound nucleic acid sequences. When $[C_f]$ is low, the rate of hybridisation is inversely proportional to complexity over a 400-fold range, indicating that the reaction is controlled by the nucleation step. However, when the hybridisation reaction is limited by diffusion of the probe to the filter, that is when $[C_f]$ is high, the rate of reaction is independent of complexity (9).

3.3 **Molecular Weight of the Probe**

For DNA-DNA hybridisation in solution, the rate is directly proportional to the square root of the molecular weight of the nucleic acid (12) and this also describes the reassociation of the probe in solution during filter hybridisation (9). However, the effect of the molecular weight of the probe on the rate of hybridisation to filter-bound sequences contrasts sharply with that found in solution. Two situations can occur. When $[C_f]$ is low compared with $[C_s]$, that is, a nucleation-limited reaction, the rate of hybridisation is independent of the molecular weight (9,10). When $[C_f]$ is high compared with $[C_s]$, that is, diffusion-limited filter hybridisation, the rate of hybridisation is inversely proportional to the molecular weight of the probe, but there are insufficient data for an exact relationship to be formulated. The observed rate of hybridisation is significantly depressed by an increase in the molecular weight of a single-stranded probe (which is not capable of reassociation). This effect is even more pronounced when a double-stranded probe is used. This is because the combined effects of a lower rate of hybridisation and the increased rate of reassociation, which accompanies an increase in molecular weight of the sequences in solution, result in lower observed rates of hybridisation and a reduced final yield of hybrid. The difference in dependence on molecular weight of the two types of filter hybridisation is not understood.

Exhibit B

DNA PROBES

Background • Applications • Procedures

SECOND EDITION

George H. Keller Mark M. Manak

M
stockton
press

Published in the United States and Canada by
STOCKTON PRESS, 1993
49 West 24th St., New York, N.Y. 10010, USA

ISBN 1-56159-102-5

First published in the United Kingdom by
MACMILLAN PUBLISHERS LTD, 1993
Distributed by Globe Book Services Ltd
Brunel Road, Houndmills,
Basingstoke, Hants RG21 2XS, England

A catalogue record for this book
is available from The British Library.

ISBN 0-333-57384-6

Printed in Great Britain

1. Probe length should be between 18-50 nucleotides. Longer probes will result in longer hybridization times and low synthesis yields; shorter probes will lack specificity.
2. Base composition should be 40-60% G-C. Non-specific hybridization may increase for G-C ratios outside of this range.
3. Be certain that no intra-probe complementary regions are present. These may result in the formation of "hairpin" structures which will inhibit hybridization of the probe.
4. Avoid sequences containing long stretches (more than four) of a single base (i.e., -GGGGG-).
5. Once a sequence meeting the above criteria has been identified, computerized sequence analysis is highly recommended. The probe sequence should be compared with the sequence region or genome that it was derived from, as well as to the reverse complement of the region. If homologies to non-target regions greater than 70% or 8 or more nucleotides in a row are found, that probe sequence should not be used.

Following these guidelines does not guarantee that a useful oligonucleotide probe will result, but it greatly enhances the chance of success. The final test is to synthesize, label and hybridize the probe to specific and non-specific target nucleic acids over a range of hybridization temperatures.

HYBRIDIZATION RATE

Traditional analysis of hybridization rate was based on DNA reassociation studies. Under those conditions, probe and target strands were present in solution and at equal concentrations. Modern hybridization experiments are often conducted in probe excess whether they involve solution hybridization or solid-phase hybridization to immobilized targets (filters, beads). However, in the case of target capture hybridization, the probe is immobilized and the target nucleic acid is present in excess. In fact, since the solid-phase targets or probes are not in solution, realistic concentrations cannot be calculated for these target nucleic acid species. For these reasons, the traditional second-order rate formulas usually cited for hybridization reactions will not be discussed, only first-order kinetic relationships will be described.

In probe-excess situations, the hybridization rate is mainly dependent upon probe length (complexity) and probe concentration. The first order formulae presented below are descriptive of single-stranded probes present in excess over target sequences. Double-stranded probes exhibit similar kinetics at short (1-4 hours) hybridization times, but not at longer times because of reassociation of the probe, which decreases the available probe concentration. Equation [1.1] (Meinkoth and Wahl, 1984) can be used to estimate the time required to hybridize half of the probe to its immobilized target sequences:

$$t_{1/2} = \frac{\ln 2}{k\,C} \qquad [1.1]$$

k = rate constant for hybrid formation (mol liter/nucleotides s)
C = probe concentration in solution (mol of probe molecules/liter)

The rate constant, k, is dependent upon probe length (L), probe complexity (N), temperature, ionic strength, viscosity and pH. L = N for probes which contain no repeated sequences. For example, for a 40-mer that contains two copies of a 20 nucleotide sequence, L = 40 and N = 20. The relationship of k to these variables is:

$$k = \frac{k_n L^{0.5}}{N} \qquad [1.2]$$

k_n is the nucleation constant and is 3.5×10^5 for Na^+ concentrations of 0.4-1.0 M, pH values of 5-9 and hybridization temperatures 25°C below the T_m of the probe-target hybrids (Wetmur and Davidson, 1968). To calculate the rate, in seconds, for hybridization of half of the probe to its target, equations [1.1] and [1.2] can be combined to give:

$$t_{1/2} = \frac{N \ln 2}{[1.3]3.5 \times 10^5 (L^{0.5})\,C} \qquad [1.3]$$

For a probe 500 nucleotides in length, the numbers would be:

$$t_{1/2} = \frac{500\ (0.693)}{3.5 \times 10^5\ (22)\ 6 \times 10^{-10}} = 75{,}000 \text{ s or } 20 \text{ hours}$$

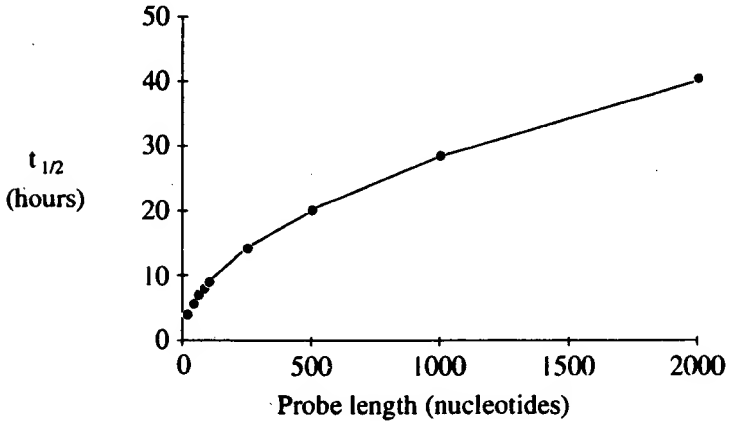


Figure 1.4 Effect of Probe Length on Hybridization Rate at Constant Molar Probe Concentration. Data points were calculated using equation [1.3] and a constant *molar* probe concentration (6.1×10^{-10} M) equivalent to a 500 nucleotide probe at 100 ng/ml.

Figure 1.4 graphically shows the relationship between $t_{1/2}$ and probe length for a range of probe sizes and a constant *molar* concentration of probe. The very long times for probes longer than 500 nucleotides illustrate the importance of using relatively short probes and the usefulness of hybridization accelerators, since hybridization times exceeding 18 hours are impractical. These times must be considered estimates because the increase in rate with probe length predicted in equation [1.3] is probably not always valid. This is due to the wide range of possible probe sizes (the factor L may not adequately compensate for probes >1 kb because of diffusion and viscosity effects), as well as the fact that all of the immobilized target may not be available for hybridization.

PROBE CONCENTRATION

In general, hybridization rate increases with probe concentration. Also, within narrow limits, sensitivity increases with increasing probe concentration. In our experience (with long probes), the concentration limits for sensitivity are approximately 5-100 ng/ml for ^{32}P-labeled probes in filter hybridizations, 25-1,000 ng/ml for non-radioactively-labeled probes in filter hybridizations and 0.5-5.0 µg/ml for *in situ* hybridizations with either type of label. The concentration limit is not determined by any inherent physical property of nucleic acid probes, but by the type of label and non-specific binding properties of the immobilization medium involved.

STRINGENCY

Factors that affect the stability of hybrids determine the stringency of the hybridization conditions. Since hybridization occurs most readily at 25°C below the T_m of the hybrids, the calculation of T_m is a necessary first step. Equation [1.4] illustrates the relationship between factors which determine T_m. The stringency can be adjusted as required by changing salt concentration, temperature or formamide concentration. For DNA:DNA hybridization using probes of more than 20 nucleotides:

$$T_m = \frac{81.5°C + 16.6 \log M + 0.41(\%G+C) - 500}{n - 0.61(\%\text{formamide})} \qquad [1.4]$$

M = $[Na^+]$ in mol/liter
n = length of shortest chain in duplex

Thus for a probe of 500 nucleotides, containing 42% G + C, in 5x SSC (0.75 M Na^+) and 50% formamide:

$$T_m = 81.5 + (-2.07) + 17.22 - 1 - (30.5) = 65°C$$

$$T_{hyb} = 65°C - 25°C = 40°C$$

Exhibit C

Hybrid Review

Critical Reviews in Biochemistry and Molecular Biology, 26(3/4):227–259 (1991)

DNA Probes: Applications of the Principles of Nucleic Acid Hybridization

James G. Wetmur

Department of Microbiology, Mount Sinai School of Medicine, New York, New York

Referee: Jacques Fresco, Dept. of Biochemical Sciences, Lewis Thomas Laboratory, Princeton, NJ 08544

ABSTRACT: Nucleic acid hybridization with a labeled probe is the only practical way to detect a complementary target sequence in a complex nucleic acid mixture. The first section of this article covers quantitative aspects of nucleic acid hybridization thermodynamics and kinetics. The probes considered are oligonucleotides or polynucleotides, DNA or RNA, single- or double-stranded, and natural or modified, either in the nucleotide bases or in the backbone. The hybridization products are duplexes or triplexes formed with targets in solution or on solid supports. Additional topics include hybridization acceleration and reactions involving branch migration. The second section deals with synthesis or biosynthesis and detection of labeled probes, with a discussion of their sensitivity and specificity limits. Direct labeling is illustrated with radioactive probes. The discussion of indirect labels begins with biotinylated probes as prototypes. Reporter groups considered include radioactive, fluorescent, and chemiluminescent nucleotides, as well as enzymes with colorimetric, fluorescent, and luminescent substrates.

KEY WORDS: Nucleic acid hybridization, thermodynamics, kinetics, single-stranded DNA branch migration, labeled probes.

I. INTRODUCTION

Nucleic acid hybridization with a labeled probe is the only practical way to detect the presence of a complementary target sequence in a complex nucleic acid mixture. The goal of this article is to develop a unified approach to all types of nucleic acid hybridization.

Quantitative aspects, thermodynamics, and kinetics of hybridization are explained below in Section II. Three temperatures are distinguished and calculated for (1) melting of polynucleotides, (2) melting of oligonucleotides, and (3) dissociation of oligonucleotide probes from targets immobilized on solid supports. The hybrids considered include DNA duplexes and triplexes, RNA duplexes, and RNA-DNA hybrids, including in each case structures with nucleotides modified in the bases or in the backbones. Hybridization rates are calculated for the same probes reacting with targets in solution and on solid supports. Additional topics include acceleration of hybridization using inert polymers, emulsions, or proteins as well as quantitative aspects of reactions involving single-stranded DNA branch migration, including D-loop formation, R-loop formation, displacement reactions, and branch capture reactions.

The synthesis or biosynthesis and the detection of labeled probes are considered below in Section III, with a discussion of their sensitivity and specificity limits. Direct labeling is illustrated with radioactive probes, including end labels and internal labels. Indirect labels are detected after they bind proteins or other molecules that act directly or indirectly as reporters. The discussion of indirect labels begins with biotinylated probes as prototypes. Reporter groups include radioactive, fluorescent, and chem-

1040-9238/91/$.50

4,5′,8-trimethylpsoralen is attached to a radiolabeled oligonucleotide, a duplex with a long complementary polynucleotide may be trapped by photocrosslinking and T_m may be determined.[61,62] The partial duplex product may be detected following separation by alkaline agarose gel electrophoresis[63] or other methods. This photocrosslinking method extends the temperature range where oligonucleotide binding may be investigated, thus permitting more complete measurements of the effect of polynucleotide secondary structure on oligonucleotide binding.[61,62] Methylphosphonate oligonucleotides may be synthesized with attached 4′-(aminoalkyl)-4,5′,8-trimethylpsoralen,[64] for photocrosslinking, as well as with attached EDTA,[65] which, when activated by Fe^{++} + DTT, will cleave the target, acting as an artificial restriction enzyme.

b. Triple Helices (Triplexes)

Many recent studies of the hybridization of oligonucleotides with native DNA at sites permitting triple helix (triplex) formation have been carried out to explore the potential of these oligonucleotides or their derivatives as artificial restriction endonucleases or antisense oligonucleotides. Triplexes have been formed *in vitro* at pH below 7 between polypurine-polypyrimidine stretches in DNA and polypyrimidine oligodeoxynucleotides with the sequence parallel to the purine-rich strand in the DNA. Systematic studies have been directed toward several variables:[66] pH, salt concentration, % G+C, and mismatches. The thermodynamic parameters for triplex formation have been determined.[67] It is interesting to note that the enthalpy per base pair associated with forming a triplex from an oligonucleotide and a duplex was only one third that for forming a duplex from oligonucleotides. The triplets contained Hoogsteen base pairs between the T in the oligonucleotide and the A in DNA, and between protonated dC (dC^+) in the oligonucleotide and dG in the DNA. Substitution of either BrdU for thymidine or MedC for dC, or substitution of both, increased the stability of the triplexes at pH 7.0, or 7.4, respectively.[68] A construct containing a 3′-3′ bridge between opposite sense oligonucleotide segments permitted triplex formation at adjacent polypurine and polypyrimidine sequences on a duplex DNA strand.[69] More recently, a purine·purine·pyrimidine triple helix motif (T may substitute for A) has been described with antiparallel polypurine strands, increasing the range of potential targets.[70]

Triplex-forming oligonucleotides containing EDTA·Fe^{++} have been activated to produce specific, although very incomplete, cleavage of the DNA.[71] To date, the greatest specificity demonstrated has been a single cleavage of a yeast chromosome into which a 20 nucleotide target had been engineered.[72]

Intercalating agents have been added to increase binding affinity by short triplexes. Triplex-forming oligo-dT oligonucleotides containing both 5′-EDTA·Fe^{++} and 3′-acridine have been investigated.[73] Sequence specificity of binding was demonstrated by footprinting for an oligonucleotide with acridine attached to the 5′-end,[74] although the specificity test was quite limited compared with that demonstrated above for a 20-mer oligonucleotide and yeast DNA.[72]

In addition to EDTA·Fe^{++}, other covalent additions used to mediate detection have included ellipticine,[75] which permits photocleavage, *p*-azidophenacyl,[76] which permits photocrosslinking, a N-bromoacetyl group,[77] which mediates alkylation and staphylococcal nuclease.[78]

B. Hybridization

We are using the term hybridization to include all intermolecular duplex formation by complementary nucleic acid strands, including not only RNA-DNA hybridization, but also DNA renaturation or reassociation. Hybridization reactions are illustrated in Figures 1b and 2b.

1. Hybridization with Polynucleotides

The fundamental studies of hybridization, including the second-order kinetics[79] and the dependence on polynucleotide length[80] and information content or complexity,[79-81] were completed in 1968. The theoretical and practical aspects of hybridization have not changed appreciably since

the subject was reviewed in 1974.[82,83] A comprehensive book on hybridization was published in 1985.[84] Thus, what appears below will be a short summary of the subject, initially directed toward quantitative interpretation of solution hybridization reactions.

Because the individual polynucleotide strands often do not contain the same sequences, the concentration term used for hybridization of polynucleotides is molar nucleotide residue concentration (C_o), not strand concentration (C), as is used for both oligonucleotide denaturation or hybridization.

The rate constant, (k_2), in $M^{-1}\ s^{-1}$, for hybridization reactions is given by

$$k_2 = \frac{k_N' \sqrt{L_s}}{N} \qquad (3)$$

where L_s is the length of the shortest strand participating in duplex formation, N is the complexity or the total number of base pairs present in non-repeating sequences, and k_N' is the nucleation rate constant. The inverse dependence of k_2 on N results from mass action, where at constant C_o increasing N means a lower concentration of any particular sequence.

Because the yield of base pairs for a given nucleation increases as L_s, the dependence of k_2 on $\sqrt{L_s}$ implies that fewer nucleation sites are available for reaction as the molecules get longer. Both experiments[12,80,85] and excluded volume theory[80,85,86] agree that circular permutation does not affect the yield of base pairs formed.

Hybridization reactions are second order if the concentrations of the complementary strands are equal and pseudo-first order when the concentration of one strand is in excess. In Equations 4a and 4c, f_{ss} is the fraction of potential duplex remaining single-stranded at time *(t)*.

Second order:

$$\frac{1}{f_{ss}} = k_2 \cdot \frac{C_o}{2} \cdot t + 1 \qquad (4a)$$

for which the half time (seconds) is

$$t_{1/2} = \frac{2}{k_2 \cdot C_o} \qquad (4b)$$

Pseudo-first order:

$$f_{ss} = e^{-k_2 \cdot C_o \cdot t} \qquad (4c)$$

for which the half time (seconds) is:

$$t_{1/2} = \frac{\ln 2}{k_2 \cdot C_o} \qquad (4d)$$

Equations 4a and 4c are used for plotting rate data. For the second-order reaction, a plot of $1/f_{ss}$ vs. *t*, in seconds, gives a slope equal to $k_2 \cdot C_o/2$. For the first-order reaction, a plot of $\ln (1/f_{ss})$ vs. *t* gives a slope equal to $k_2 \cdot C_o$.

Many hybridization reactions are carried out with *excess target* DNA or RNA, using tracer concentrations of labeled probe compared with the concentration of complementary sequences present in the target. Equations 4a and 4b describe reactions between single- or double-stranded probes and double-stranded targets in solution. For all excess target reactions, C_o is the total (mostly target) nucleotide concentration.

Example 4b

Let $G_0 = 1.5 \cdot 10^{-4} M$ (the same as 50 μg/ml and, for native DNA, absorbance 260 nm = 1). Let $[Na^+] = 1$ such that $k_N' = 3.5 \cdot 10^5 M^{-1}\ sec^{-1}$. Digest the DNA with a restriction enzyme with a 4-base recognition sequence (frequency = $1/4^4$). The weight average length (nucleotides) is twice the number average (strands), giving $L_s = 512$. For *Escherichia coli* DNA: $N = 4.2 \cdot 10^6$.

$$k_2 = \frac{k_N' \sqrt{L_s}}{N} = 3.5 \cdot 10^5 \cdot \sqrt{512}/N$$
$$= 1.9 M^{-1}\ sec^{-1}$$

$$\text{and } t^{1/2} = \frac{2}{k^2 \cdot C_o} = 2/(1.9 \cdot 1.5 \cdot 10^{-4})$$

$$= 7000 \text{ seconds} = 1.95 \text{ hours}$$

Single-copy target DNA:

increase the stringency or fidelity of hybridization with excess target. Alternately, with excess probe, reactions may be carried out at T_m^∞-25°C, leading to mismatched products, and these structures may be subsequently eliminated by melting (e.g., washing at low salt).

Because k_N' is a very strong function of ionic strength at salt concentrations below 0.2*M*, hybridization reactions should not be carried out in low salt. In the range of interest ($0.25 \leq [Na^+] \leq 4.0$), the dependence of k_N' on salt concentration is given by Equation 6.[88]

$$k_N' = \{4.35 \cdot \log_{10}[Na^+] + 3.5\} \cdot 10^5 \quad (6)$$
$$\text{for } 0.2 \leq [Na^+] \leq 4.0$$

Hybridization reactions may be carried out in denaturing solvents while maintaining high ionic strength. In these cases, k_N' is reduced if the denaturing solvent has a higher viscosity than k_N' in 1.0*M* NaCl at 70°C. Examples include DNA hybridization in 2.4*M* Et_4NCl or other tetraalkylammonium salt solutions[88] and formamide,[89] where k_N' decreases 1.1% for each 1% addition of formamide. RNA hybridization may also be carried out in tetraalkylammonium salt solutions. In Et_4NCl, k_N' for RNA hybridization is almost an order of magnitude lower than k_N' for RNA hybridization. The origin of this difference is unknown.

In most solution hybridization reactions, the rate of diffusion of the probe is sufficient to assure reaction homogeneity. However, in reactions with trace probe concentrations, stirring may be necessary to assure homogeneity during the latter part of the reaction.[90]

Mismatches of up to 10%, although easily measured as a depression in T_m^∞, have essentially no effect on hybridization rates at the temperature of the maximum rate and may be ignored.[29] Thus, most of the modifications used for labeling probes described below in Section II, even if they significantly decrease T_m^∞ and T_m, have little or no effect on k_2. As mismatching is increased to 20 and to 30%, where T_m^∞ is nearing the normally optimum hybridization temperature, k_N' falls to half and then to zero.

Subtractive hybridization, based on duplex structure with[91] or without[92] an affinity label, has been used to deplete one complex nucleic acid mixture, such as human DNA, of all sequences present in a second complex mixture, such as human DNA containing a deletion mutation. Subtractive hybridization with mammalian DNAs is facilitated by multiple copies of the desired product in the second mixture,[92] as well as by using one of the methods of acceleration of hybridization described below in Section II.B.4.[92] Recently, subtractive hybridization has been carried out without multiple copies or use of an acceleration method by using PCR to amplify the products following subtraction.[93]

Possible DNA hybridization reactions in agarose gels include reassociation of bacterial DNA,[94] permitting detection of restriction fragment length polymorphisms, and mammalian DNA,[95-97] permitting detection of amplified sequences. The method may be extended to permit differential cloning of restriction fragment length polymorphisms in single copy mammalian DNA[98] using another of the hybridization acceleration methods described below in Section II.B.4. As a first approximation, the hybridization rates in gels may be predicted by assuming the same parameters in the k_2 calculation for hybridization in the appropriate solvent and the same C_o as was present in the well prior to electrophoresis.

2. Hybridization with Oligonucleotides

a. Duplexes

Hybridization rates with oligonucleotides are also described by Equations 3 and 4a through d. The reactions are usually reported in terms of molar concentration of oligonucleotide (C), and not the nucleotide concentration (C_o). C_o is C *multiplied by the length of the oligonucleotide.*

The definition of complexity remains the same for oligonucleotides and polynucleotides. It is important to note that the increased complexity of mixed oligonucleotide probes used for screening cDNA libraries[99] increases the half-time for hybridization proportionally. The $\sqrt{L_s}$ dependence of k_2 has not been strictly verified for oligonucleotides, although estimates based on Equation 3 are remarkably accurate.[15] Equation 3 may underestimate k_N' for the shortest oligonucleotides where the availability of nucleation

sites may not decrease with increasing L_s. The k'_N for poly(rA) + poly(rU)[14] is $8 \cdot 10^4 M^{-1}\ s^{-1}$ after correction from 0.4 to 1.0*M* Na^+.

Compare the forward rate constant[100] for self-association of $r(A_7U_7)$ at $100 \cdot 10^4 M^{-1}_{strands}\ s^{-1}$: dividing by the length of 14 to get units of nucleotides and correcting the four-fold increase in rate for self-complementary reactions, $k_2 = \{100/(4 \cdot 14)\} \cdot 10^4 M^{-1}\ s^{-1}$. Equation 3 with $L_s = N = 14$ for $r(A_7U_7)$ predicts: $k'_N = k_2 \cdot \sqrt{L_s} = \{100 \cdot \sqrt{14}/(4 \cdot 14)\} \cdot 10^4 M^{-1}\ s^{-1} = 7 \cdot 10^4 M^{-1}\ s^{-1}$.

In this case, the same result is found for both oligonucleotides and polynucleotides. On the other hand, a similar calculation using a forward rate constant[38] of $4 \cdot 10^4 M^{-1}\ s^{-1}$ for d(GGAATTCC) gives $k'_N = 35 \cdot 10^4 M^{-1}\ s^{-1}$ in 0.25*M* Na^+, which is about fourfold higher than predicted by Equation 6.

With polynucleotides, k'_N is affected by temperature, ionic strength, and viscosity. Just as with polynucleotides, oligonucleotide RNA hybridization rates are less than DNA hybridization rates, which are nearly the same as RNA-DNA hybridization, depending on % G+C.[101]

The temperature dependence of the rate of oligonucleotide hybridization is simpler than the bell-shaped dependence of k_2 on temperature. With polynucleotides near T_m^∞, k_2 decreases with increasing temperature as base pair formation rates at the end of duplex regions approach base pair dissociation rates. Hybridization with oligonucleotides usually takes place far below T_m^∞. In fact, there is little temperature dependence of k_2 for oligonucleotide-oligonucleotide reactions.[38,101] Reaction rates between oligonucleotides and polynucleotides at temperatures above T_m but still well below T_m^∞, determined using psoralen photocrosslinking, indicate a temperature dependence attributable to intramolecular pairing of the polynucleotide strand.[61,62] The effect of solvent viscosity on oligonucleotide hybridization rates has not been studied adequately.

The effect of ionic strength on T_m is the same for oligonucleotides and polynucleotides. For example, the slope of a plot of (log k_2) vs. ($-$log $[Na^+]$) is 1.3 for DNA[80] in the interval $0.1 \leq [Na^+] \leq 0.7$ and 1.4 for self-complementary d(GGAATTCC)[38] in the interval $0.05 \leq [Na^+] \leq 0.3$. Because the $[Na^+]$ dependence of T_m decreases for both oligonucleotides and polynucleotides at high $[Na^+]$, Equation 6 may be used for all hybridizations in the practical interval indicated.

One of the steps in PCR with thermostable polymerases is an annealing step.[4] PCR primers are typically present at C = 1 μM and are commonly 20 nucleotides long. Given that k'_N is $5 \cdot 10^4 M^{-1}\ s^{-1}$ in PCR buffer (equivalent to 0.20*M* NaCl), using Equations 3 and 4d, we predict $t_{1/2} = 3$. By the time the annealing step temperature is reached, annealing is completed. At C = 0.01 μM used for oligonucleotide ligation assays, although $t_{1/2}$ is increased to 300 due to lower C and may be further increased in formamide-containing solvents. The $t_{1/2}$ is still much less than the ligation time and is not a variable in the assay.

b. Triple Helices (Triplexes)

A recent study of the kinetics of triplex formation, employing a restriction endonuclease protection assay, has led to the surprising result that triplex formation is quite slow.[102] All of the conditions studied, with 0.07 or 0.6*M* NaCl, corresponded to high salt conditions because of the presence of 400 μM spermine and 20 m*M* Mg^{++}. The forward strand rate constant observed at 37°C was approximately $1800 M^{-1}\ s^{-1}$ or, dividing by the length of 21, $k_2 = 86 M^{-1}\ s^{-1}$. Even after correcting k'_N to $2 \cdot 10^5 M^{-1}\ s^{-1}$ at 37°, k_2 from Equation 3 for duplex formation was 500 times greater than k_2 observed for triplex formation.

3. Reactions on Solid Supports

a. Blots

Most hybridization reactions on solid supports occur with target nucleic acids noncovalently or covalently linked to nitrocellulose or to a nylon-based (or other) membrane. If a probe is hybridized to a membrane replica of a library of plasmid-containing bacteria, or of bacteriophages, following alkaline lysis *in situ* and immobilization of the DNA on the membrane, the procedures are called colony or plaque hybridizations, respectively. A similar screening procedure may be used with yeast artificial chro-

mosome clones.[103] If the target DNA or RNA is applied directly to the membrane, the probe hybridization reaction is a dot blot, a slot blot, or simply a filter hybridization. Sandwich hybridizations may be carried out using two adjacent, nonoverlapping probes.[104,105]

If the target is DNA that has been separated by gel electrophoresis prior to transfer (blotting) to the solid support, the reaction is called Southern hybridization.[3] Southern hybridization may be carried out with large DNAs separated by pulsed field gel electrophoresis using several types of apparatuses.[106-110] Southern hybridization may also be carried out with small DNA fragments that are obtained from sequencing gels, allowing multiplex,[111] as well as direct[112] sequencing of complex DNAs. If the target is RNA that has been separated by gel electrophoresis, the reaction is called Northern hybridization.

The fundamental aspects of hybridization on solid supports have not changed appreciably since the topic was the subject of a review in 1984[113] and a chapter in a comprehensive book on hybridization published in 1985,[84] although additional practical details may be found in recently updated cloning manuals.[114,115] All of the results of variables affecting melting temperatures, detailed above for solution hybridization, are applicable for hybridization on solid supports. The effects of temperature, ionic strength, denaturing solvents, and solvent viscosity on k'_N, and the lack of effect of base composition or limited mismatching on k'_N, are the same in solution and on solid supports.

Equation 3 needs to be reconsidered. The effect of complexity on k_2 is the same for hybridization in solution and on solid supports, but the effect of length is more complicated. When the probe is an oligonucleotide, Equation 3 is found to be valid. The same rates of hybridization are also found for 50-nucleotide probes simultaneously hybridizing with 150-nucleotide DNA in solution and tethered once to latex beads.[116] The newly described method for introduction of a residue during phosphoramidite synthesis, which may be used later with a variety of tethers,[117] should permit a more thorough study of surface effects on hybridization with bound molecules.

More is known about filter hybridization. In this case, the target nucleic acids are tethered to the solid support at many sites. Each short domain between tethers acts as an independent target. Thus, as the length of a polynucleotide RNA probe is increased, the rate of hybridization to immobilzed DNA is unchanged following RNase digestion and washing of the hybrids on the membranes.[118] At $C_o = 2.75\ \mu M$, $t_{1/2} = 4800$ s for ϕX174 (N = 5386) RNA-DNA hybridization in high salt, 50% formamide. According to Equation 4d, $k_2 = 53$. According to Equation 3, $k'_N \cdot \sqrt{L_s} = 5380 \cdot 53 = 2.9 \cdot 10^5 M^{-1}\ s^{-1}$, which is 2 to 4 times the value of k'_N for RNA-DNA solution hybridization previously reported in this solvent. Apparently the first approximation to calculation of predicted k_2 values is to use Equation 3 with L_s set equal to 10, no matter what the actual length of the probe.

Equations 4a through 4d also need to be reconsidered. For those extremely rare occasions where the probe is at trace concentrations compared with the complementary target sequences on the membrane, Equations 4c and 4d are applicable for both single- and double-stranded probes reacting with both single- and double-stranded target DNA or RNA, because the target DNA or RNA cannot hybridize with itself. The complexity and concentration of the target still govern the reaction, although this situation may be complicated if the probe is so dilute and the target is so concentrated that hybridization depletes the probe in the vicinity of the membrane. Then the reaction becomes dependent upon diffusion of the probe to the target.

The more common situation is excess probe. If the probe is single-stranded, Equations 4c and 4d again hold, but the complexity and concentration of the *probe* govern the reaction. If the probe is double-stranded, the problem of competing reactions occurs. The probe should not be longer than the effective length of the target strands, or the rate of probe-probe hybridization will exceed the rate of probe-target hybridization, and the kinetics will become much more complicated. A length mismatch does not affect the initial rate of hybridization, but becomes important at later times. Because such a length mis-